Using Choice Experiments to Estimate Consumer Valuation: the Role of Experimental Design and Attribute Information Load

Zhifeng Gao[1], Xiaohua Yu[2] and Lisa O. House[3]

Authors are:
1: Assistant Research Scientist
3: Professor
both at
Food and Resource Economics Department
University of Florida

2: Junior Professor
Courant Research Centre -PEG
University of Goettingen

Email:
Zhifeng Gao: zfgao@ufl.edu
Xiaohua Yu: xyu@uni-goettingen.de
Lisa O. House: lahouse@ufl.edu

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association Annual Meeting, Milwaukee, WI, July 26-28, 2009

Introduction

Choice experiments (CE) or choice-based conjoint analysis have been widely used to elicit consumer valuation of nonmarket goods (i.e. Adamowicz et al. 1998; Boyle et al. 2001; Colombo et al. 2007) and marketable goods with novel attributes (i.e. Alfnes 2004; Darby et al. 2008; Tonsor et al. 2005). The increasing popularity of CE types of surveys is partially in response to recognized problems of contingent valuation by the NOAA panel in 1990 (Hausman 1993), and its ability of easily identifying the trade-off among different product attributes relative to other approaches. However, several issues with use of CE remain unresolved. The major challenge is how to design statistically efficient experiments to provide enough information for accurately eliciting consumer preferences, and at the same time, to make the length of choice experiments reasonable such that cognitive burdens on survey participants are minimized.

Although different types of design strategies have been developed (see Louviere, Hensher and Swait (2000) for various design strategies), no general agreement has been reached on what is the "best design" of choice experiments. Each design approach has its advantages in capturing certain types of effects and there is no superior design for all purposes (Chzan and Orme 2000). D-optimal design with correct priori information generates more accurate valuation of products or services (Carlsson and Martinsson 2003; Ferrini and Scarpa 2007), however, if high quality prior information is not available, the shift design is the most promising (Ferrini and Scarpa 2007). Lusk and Norwood (2005) compared six design strategies regarding their performance in WTP estimates and demonstrated that random designs are the best. In addition, designs incorporate attribute interactions results in more precise valuation estimates than main effects only design (Lusk and Norwood 2005). Previous research recognizes that the criterion of a good choice experiment design is not only in the statistical efficiency, but also in the cost associated with the choice complexity. This is because more statistically efficient design is always accompanied by a higher level of complexity which results in heavier cognitive burden on respondents. However, existing literature using simulations has only compared different design strategies with fixed design dimensionality (e.g., number of alternatives, number of attributes and number of attribute levels) which is highly correlated with the complexity of choice experiments. For instance, Carlsson and Martinsson (2003) constructed pair-wise choice experiments with 4 attributes, three of them having 3 levels, and one having 2 levels; Lusk and Norwood (2005) designed choice experiments in which each choice set consisted of 3 alternatives, and each alternative with four 3-level attributes. Ferrini and Scarpa's (2007) conducted a simulation based on pair-wise choice experiments with 4 attribute, each having 3 levels.

Several studies have demonstrated that design dimensionality, especially the number of product attributes included in choice experiments, affects consumer preference and valuation (Islam, Louviere, and Burke 2007, Hensher 2006, Gao and Schroeder 2009). Lack of research on the performance of different experimental designs under various attribute information, implied by the number of product attributes, hinders our ability to infer efficiency of different choice experiment design strategies. In addition, it is also unclear if the impacts of the number of product attributes in choice experiments result from its effects on respondent cognitive ability of processing information or the statistical property of choice experiments when the number of attributes increases. The purposes of this article are to 1) determine the performance of different choice experimental designs on welfare estimates such as willingness-to-pay (WTP) under different attribute information loads; and 2) investigate the effects of information loads on consumer valuation in simulation scenarios. We believe that more appropriate design strategies can be made base on the attribute information determined by researchers before implementation of choice experiments. In addition, identifying the impacts of dimensionality on the statistical property of choice experiment helps enhance our understanding of the effects of dimensionality on respondent valuation.

Choice Experiment and Good Experimental Design

In a choice experiment, predetermined attributes that are believed to have the largest impacts on consumer choice decisions comprise a series of alternatives (profiles or choices). Two or more alternatives are used to form a choice set, and a sequence of choice sets composes a choice experiment. Respondents are asked to choose one alternative from each choice set in an experiment. Based on the random utility theory, a consumer will choose an alternative from each choice set to maximize her/his utility. Consumer preferences for products and product attributes can be elicited from their sequent choice decisions. However, in most cases, enumerating all combinations of product attributes is not feasible because the number of combinations of product attributes is huge, resulting in very large number of choice sets. Too many choice sets hinder consumer ability to make rational choice decisions in a short time. Therefore, one of the major challenges of conducting a choice experiment is to design choice experiments simple for respondents to make efficient choice decisions while at the same time provide enough information for researchers to accurately elicit consumer preferences.

Various experimental designs have been discussed and used by researchers, including orthogonal main effects design, D-optimal design, fractional factorial design and shifted (or cycled) design, etc. There is a general agreement that the choice experiment generated from the design should result in the minimum variance of coefficient estimates. In a linear model, the variance of coefficient estimates does not depend on the true parameters, such that

$Var(\beta) = \sigma^2 (X'X)^{-1}$, where X is the design matrix and σ^2 is the variance of the random error in the linear model. Therefore, the D-optimal design which maximizes the D-efficiency $100 \cdot \frac{1}{N\,|(X'X)^{-1}|^{1/(A \cdot C)}}$, should be the best design method in linear models, where N is the number of observation in the design, A is the number of attributes and C is the number of attribute levels. The balanced orthogonal design automatically results in a design with 100% D-efficiency because $(X'X)^{-1} = \frac{1}{N} \cdot I$ where I is an identity matrix. A design deviated from unbalanced but orthogonal design will have a D-efficiency less than 100, with the minimum efficiency being zero (Kuhfeld, Tobias and Garrat 1994). The algorisms that maximize D-efficiency of a design are readily available in SAS and other software which make the design of experiments of linear models simple. However, good designs based on the standards in a linear model may not hold for choice experiments. The models used in choice experiments are usually nonlinear, and the variance of parameter estimates not only depends on the design matrix, but also on the true parameter in the models.

The variance matrix of parameter estimates from choice experiments depends on the assumption of the random component in consumer random utility function $U_j = \beta' x_j + \varepsilon_j$, where U_j is consumer utility of consuming product j, β is a vector of parameters, x_j is a vector of attributes for alternative j and ε_j is the random component with a certain type of probability distribution. Consumers will choose an alternative from a choice set to maximize her/his utility and the probability of choosing alternative j is $P_j = \text{Prob}(\beta' x_j + \varepsilon_j > \beta' x_i + \varepsilon_i,\ \forall i \neq j)$. In a multinomial logit model (MNL), a prevailing model used to estimate consumer preferences in choice experiments, ε_j has identical independent Gumbel distribution. The probability that a consumer chooses alternative *j* from a choice set with *n* choices is $P_j = \frac{e^{\beta' x_j}}{\sum_{i=1}^{n} e^{\beta' x_i}}$. The maximum likelihood estimator is consistent and asymptotically normal distributed with mean β, and the asymptotic covariance matrix $\Omega = (Z'PZ)^{-1} = [\sum_{n=1}^{M} \sum_{j=1}^{n} z'_{jn} P_{in} z_{jn}]^{-1}$, where $z_{jn} = x_{jn} - \sum_{i=1}^{n} x_{in} P_{in}$, and M is the number of choice sets for all respondents (McFadden 1974). The formula of the variance of the parameters in MNL implies that it is not possible to choose a design strategy that minimizes the variance of parameter estimates without knowing the true parameters in the consumer utility function. Carlsson and Martinsson (2003) and, Ferrini and Scarpa (2007) demonstrate that the choice experiment maximizing the D-efficiency $100 \cdot |\Omega^{-1}|^{1/(A \cdot C)}$ with high-quality prior information results in most accurate parameter estimates. However, in most cases, obtaining and verifying high-quality prior information of the true parameters is difficult and designs maximizing D-efficiency with low quality prior information usually leads to inferior estimation to those designs without prior information (Ferrini and Scarpa 2007).

In this article we extend Lusk and Norwood (2005) work by investigating the efficiency of different experimental designs in various information loads used to estimate consumer valuation. Lusk and Norwood (2005) compared the impacts of experiment designs on consumer valuation without prior information on consumer preferences. Like other research, their work only investigated the impacts of experiment designs with fixed product attributes. However, the amount of attribute information provided affects consumer behaviors, so that it is important to understand the performance of different experimental designs under various attribute information loads and identify sources of impacts of attribute information on consumer preferences.

Comparison of Different Experiment Designs under Various Information Loads

Comparison Scenarios

To investigate the performance of different experiment designs under various attribute information loads, we conduct several Monte Carlo simulations where the true consumer utility functions are known. Pair-wise choice experiments are used in the simulations in which respondents are assumed to choose one alternative from two choices in a choice set. With the true utility function, random components following a Gumbel distribution are added to simulate unobservable consumer preferences that may affect the choice decision. Consumer choices are simulated by comparing the utility level of the two alternatives in a choice set and then a MNL model is used to estimate the utility parameters. Consumer WTP for product attributes from MNL are compared with the true WTP, and the deviation can be used for evaluating different experimental designs.

We employ four types of choice experimental designs and evaluate their performances under four attribute information loads, three levels of sample size and two types of utility functions (table 1). Therefore the total number of simulation scenarios is $4 \times 4 \times 3 \times 2$. The choice experiment design includes randomly drawn choice sets from $3^N \times 3^N$ full factorial design (RD), where *N* indicates the number of attributes in a choice experiment; main effects only design drawn from $3^N \times 3^N$ full factorial design (ME); main effects design maximizing D-efficiency with a minimum number of choice sets drawn from $3^N \times 3^N$ full factorial design (MD); and the design pairing alternatives generated from 3^N full factorial design using bin method (RP). Those four designs are similar to those in Lusk and Norwood (2005). We exclude designs with two-way interaction effects, because the main effects design is the most commonly used design. In addition, adding two-way interactions significantly increases the choice sets in a choice experiment, especially when the attribute information load is high, thus resulting in heavy cognitive burden on survey respondents. Smaller main effects only design may be preferable even if larger designs with interactions have statistical advantages (Lusk and Norwood 2005). The numbers of attributes, each with three levels in choice experimental designs are 2, 3, 4 and 5 implying the attribute information loads from low to high. The minimum number of attributes is selected to be two, because *price* and another product attribute must be included in order to calculate the WTP estimate. The maximum number of attributes is five, because a larger number of attributes typically leads to choice experiments with more choice sets, which are difficult to administer and also quickly result in respondents' fatigue or information overload. Similar to Lusk and Norwood (2005) a continuous and a discrete true consumer utility function are assumed. The samples size measured by the total number of choice sets for all respondents are selected to be approximate 250, 500 and 1000, representing the surveys with small, middle and large sample size, respectively. The description of different simulation scenarios is provided in table 1.

The true consumer utility functions employed in our study are

(1) $V_{ij}^{c} = \alpha_j + \beta_0 \cdot P + \sum_{k=1}^{n} \beta_k \cdot x_{ijk}$ for continuous utility functions and

(2) $V_{ij}^{d} = \alpha_j + \beta_0 \cdot P + \sum_{k=1}^{n} \beta_{ka} \cdot x_{ijka} + \sum_{k=1}^{n} \beta_{kb} \cdot x_{ijkb}$ for discrete utility functions. Where α_j is alternative specified constant, β s are parameters related to product price and different attributes, *P* is product price, x_{ijk} is *kth* product attribute, x_{ijka} and x_{ijkb} are dummy variables corresponding to the levels in attribute k^{i} [1], and *n* equals 1, 2, 3 and 4 denoting the number of attributes in choice

experiments. The true parameters in continuous utility functions are $\alpha_j = 1$, $\beta_0 = -1$, $\beta_1 = 2$, $\beta_2 = 3$, $\beta_3 = 4$ and $\beta_4 = 0.1$. The true parameter of dummy variables in discrete utility functions are $\beta_{1a} = 2$, $\beta_{1b} = 1$, $\beta_{2a} = 3$, $\beta_{2b} = 2$, $\beta_{3a} = 4$, $\beta_{3b} = 3$, $\beta_{4a} = 5$, $\beta_{4b} = 4$.

Monte Carlo Simulation

For each alternative in choice experiments, the utility V_{ij}^k (*k=c, d*) is calculated using the presumed true parameters. Random variable ε_{ij} with a Gumbel distribution is drawn independently with the number of observation equal the total number of alternatives in a simulation (2 $\times$ the number of choice sets in a choice experiment $\times$ the sample size). As such, the assumptions of multinomial logit models are strictly satisfied. The consumer random utilities $U_{ij}^k = V_{ij}^k + \varepsilon_{ij}$ (*k=c, d*) are compared across alternatives in a choice set, and the alternatives with the highest random utility level are assigned a number of one to simulate consumer choices. MNL models are estimated using simulated consumer choices and the attribute levels in choice experiments. The WTP estimates for each attribute is calculated as the negative value of the ratio of attribute and price coefficients such $WTP = -\frac{\beta_k}{\beta_0}$, (*k=1-4*) for the continuous utility function and $WTP = -\frac{\beta_{kh}}{\beta_0}$, (*k=1-4, h=a, b*) for the discrete utility function. The procedure is repeated 500 times to compose a Monte Carlo simulation, resulting in 500 WTP of each attribute for each simulation scenario. The performance of different choice experimental designs with respect to each attribute evaluation is compared using mean squared error $MSE = \sum_{i=1}^{N} (WTP_i - WTP_t)^2 / N$, and absolute relative error $RSE = \frac{1}{N}\sum_{i=1}^{N} \left|(WTP_i - WTP_t) / WTP_t\right|$, where WTP_i is the simulated WTP, WTP_t is the corresponding true WTP, and N is the number of simulations. Because consumer true WTP for different attributes varies from \$0.1 to \$5, the *RSE* error provides a relative measure of the error such that the impact of different designs can be compare across attributes. The choice experiment design that results in the minimum *MSE* and *RSE* is considered here the best design. The *MSE* and *RSE* of different scenarios can be compared 1) across sample size to study the effect of sample size, 2) across different designs to investigate the performance of different designs with small, mid and large sample sizes, 3) across the number of attributes to investigate the effect of information loads on WTP estimates.

Identification of Factors Affecting WTP Estimate

Pair-wise comparisons across the sample size, design strategy and the number of attributes can help identify the impact of a single factor on WTP estimates. However, the effects of different factors may be compounded so that the pair-wise comparisons are not sufficient. Three simple models are estimated as:

(3) $WTP_i = \alpha + \beta_1 \cdot Des_i + \delta_1 \cdot Sam_i + \lambda_1 \cdot Att_i + \kappa_1 \cdot Att_i^2 + \varepsilon_i$
(4) $MS_i = \alpha + \beta_2 \cdot Des_i + \delta_2 \cdot Sam_i + \lambda_2 \cdot Att_i + \kappa_2 \cdot Att_i^2 + \varepsilon_i$ and
(5) $RS_i = \alpha + \beta_3 \cdot Des_i + \delta_3 \cdot Sam_i + \lambda_3 \cdot Att_i + \kappa_3 \cdot Att_i^2 + \varepsilon_i$

Where WTP_i is the willingness- to-pay estimate, $MS_i = (WTP_i - WTP_t)^2$ is the squared difference between the estimated WTP and the true WTP, $RS_i = |(WTP_i - WTP_t)/WTP_t|$ is the relative difference between the estimated WTP and the true WTP, $Des_i = [ME_i \ MD_i \ RP_i]'$ is a vector of dummies denoting the design strategy, $Sam_i = [M_i \ L_i]'$ is a vector of dummies denoting the sample size, Att_i is the number of attributes in the choice experiments. αs, βs, δs, λs and κs are coefficients to be estimated, and ε_i is stochastic errors. A quadratic term of Att_i is added in the models because Swait and Adamowicz (2001) demonstrate that the preference variance of consumers has a quadratic relationship with the complexity of the decision environment. Gao and Schroeder

(2009) have shown that consumer WTP has a quadratic form with the number of attributes in choice experiments. However, it is not clear whether the relationship is from changing consumer preference as a result of the complexity of decision making, or a statistical property of choice experiments with the increasing number of attributes in designs. Models (3), (4) and (5) are estimated for both continuous and discrete utility functions.

Results

WTP Estimates, MSE, RSE and Best Designs in Various Scenarios

Table 2 reports the means, *MSE* and *RSE* of simulated WTP of each attribute for the corresponding simulation scenarios when the true utility functions are continuous. For each attribute, the best experimental designs under the combinations of the number of attributes and the sample size choices are reported based on the minimum *MSE* and minimum *RSE* rules. In general, both rules result in the same best design. However, in some cases (e.g., for attribute x1, with 2 attributes in a choice experiment and a large sample size), there is divergence of the best designs between *MSE* and *RSE* rules. This may result from the *MSE* punishes extreme values in the sample more severely than the *RSE*, because of the squared difference between WTP estimates and the true WTP.

Overall, the simulated WTP is larger than the true WTP. For all designs under different information loads, increasing the sample size of choice experiments reduces the WTP estimates, which is companied with smaller *MSE* and *RSE*. The number of product attributes in choice experiments affects the performance of different designs. In general, choice experiments with 5 attributes are accompanied with higher WTP estimates, *MSE* and *RSE*. However, the impact of the number of product attributes on WTP estimated, *MSE* and *RSE* measures are not clear—it depends on the choice experiment design. For instance, in the RD design, a larger number of attributes in choice experiments is accompanied with larger WTP estimates and *MSE/RSE* as well. However, in the ME design, the WTP estimates and *MSE/RSE* are the largest when the number of attributes in choice experiments is four. Results imply it may not be the best strategy to improve WTP estimation while keeping the number of attributes at the small level, at least from a statistical perspective.

The selected best design for each attribute under different number of product attributes and sample size combinations demonstrates that there are no dominant designs. The performances of designs are highly related to the number of product attributes. For instance, with two attributes in choice experiments, ME design is the overall best design. When the number of attributes increases to 3, RP design changes to be the best design, and with the number of attributes being 4 and 5, RD design performs the best. Table 3 reports the means, *MSE* and *RSE* of simulated WTP for each attribute when the true utility functions are discrete. Results reveal the similar conclusion as from the case of the continuous utility functions. Most WTP estimates are larger than the true WTP, a larger sample size results in smaller *MSE* and *RSE* in all scenarios, and increasing the number of attributes in choice experiments leads to larger *MSE* and *RSE* in most cases. The best designs under the sample size and number of attributes combinations, however, are not consistent with the results from continuous utility function and may be attribute related. For instance, when the number of attributes in choice experiments is 2, the best design for attribute x1a is the MD design. However, for attribute x1b, the best one is the RP design. Unlike in the continuous utility function cases where the MD designs perform the worst in general, the MD designs are the best designs when the number of product attributes in choice experiments is three. When the number of product attributes is four, the ME designs become the best, and with five attributes in choice experiments, RD designs perform the best. However, it needs to point out that the best designs based on *MSE* and *RSE* may not be the "golden rule" to select the design strategy in practice. In some cases, the difference in *MSE* and *RSE* between the best and the second best design is very small, and sometimes, negligible.

Factors Affecting WTP Estimates

Results of model (3) for continuous utility functions are reported in table 4. Most of the coefficients are significantly different from zero at 5 percent significance level. The positive coefficients of design strategy indicated that RD, ME and RP designs tend to result in higher WTP estimates for attributes X1, X2 and X3, compared with the MD design. However, for the WTP for X4, those 4 design strategies are equivalent with regard to the WTP estimate. The negative signs of the sample size coefficients imply that middle and large sample size tend to have smaller WTP estimates compared with the case of small sample size. This is consistent with the fact that the WTP estimates are generally larger than true WTP. Smaller WTP estimates with middle and large sample size indicate an improvement of WTP estimate as sample sizes increase. The insignificant coefficients of Middle for WTP for attribute x2 and x3 show that increasing sample size from small to middle does not decrease the WTP. The statistical tests demonstrated that for attributes x2, x3 and x4, middle sample size is equivalent to large sample size in explaining the variation in WTP estimation. The quadratic relationship between WTP estimate and the number of attributes is statistically significant at the 5 percent level with the WTP for attribute x1. However, these effects are not significant for attribute x2, and increasing the number of attributes in choice experiments do not have significant impact on the WTP estimate of attribute x3.

Results of model (4) and (5) reported in table 4 implies that RD, ME and RP designs are significantly better than MD design based on the measures of squared or absolute relative errors. Compare to the MD design, those three designs significantly reduce the square and relative absolute errors of estimated WTP. The relatively larger coefficients for RD imply that the RD designs are the best, and the RP designs are second best. Based on the SE and RE measures, increasing sample size significantly improves the WTP estimates of all design strategies. The tests for the coefficients of the design strategy indicate that both RD and RP designs are significantly better than ME design in explaining variation in WTP estimate, SE and RE measures for the attributes x1, x2 and x3. However, the RD designs are not statistically significantly superior to the RP designs at the 5% significance level. For attribute x4, all the four design strategies are equivalent in explaining the variation in WTP estimates. For other attributes, RD, ME, and RP designs are significantly better than the MD design to explain the variations in SE and RE (5% significance level), but they are not significantly different from each other. One potential reason for this phenomenon may be that the true WTP for attribute x4 has a much smaller scale compared with the WTP for other attributes (0.1 vs. 2, 3 and 4). Another reason may be that x4 is only presented in the choice experiments with five product attributes, and performance of all design strategies are significantly affected by the number of attributes in the choice experiments. Additional models for attributes x1, x2, and x3 with the WTP estimates only from choice experiments with 5 attributes also show that all the three design strategies (RD, ME, and RP) are equivalent. Results imply that when the attribute information loads achieve a certain level (in our case, 5 attributes) none of those design strategies is dominant.

Table 5 reports the results of model (3) for a discrete consumer utility function. They provide similar but a little different information compared with the results from the continuous utility function. First of all, RD, ME and RP result in statistically significant (5% significance level) and larger WTP estimates for attributes x1a, x2a and x4a and significantly smaller WTP estimates for attribute x3a than the MD design. A large sample size results in statistically significant smaller WTP estimates than a small sample size, with the exception of the WTP estimates for attribute x2a with a middle sample size, and the attribute x4a with a large sample size. The WTP estimates from the middle and large sample sizes are not significantly different (5% significance level) for the attributes x3a and x4a. However, for the attributes x1a and x2a, a large sample size results in statistically significant and smaller WTP estimates (5% significance level). WTP estimates for the attributes x1a and x2a have statistically significant and quadratic relationship with the number of attributes in the choice experiments. For attribute x3a, increasing the number of attributes (from 4 to 5) in the choice experiment significantly reduced the WTP estimates. The design strategies of RD, ME and RP are not significantly different in explaining the variations in WTP estimates.

Results for the models (4) and (5) reported in table 4 imply that for all design strategies RD, ME and RP statistically improve the WTP estimation with smaller SE and RE than MD design. Based on the scale of the absolute value of coefficient estimates of design strategy, RD design performs the best followed by the RP design, which is the same result found in the case of continuous utility functions. Increasing sample size from small, to middle, and to large significantly improves the performance of WTP estimates. For the attribute x1a and x2a, the SE and RE measures have statistically significant and convex relationships with the number of attributes in choice experiments; both SE and RE first decrease and then increase as the number of attributes in choice experiments increases. The result implies that keeping the number of attributes at either the minimum (2 in our case) or the maximum (5 in our case) may not be the best strategy for WTP estimates. However, the positive coefficients for the # of Attributes for attribute x3a indicate that increasing the number of attributes from 4 to 5 significantly increases both SE and RE measures (5% significance level). The result is different from the case of continuous utility functions, in which increasing the number of attributes from 4 to 5 results in decreased SE and RE for the attribute x3. The statistical tests indicate that in most cases RD is a significantly better design than ME. However, the performances of RP and ME designs are not clear—results from model (4) imply that three out of the four cases (attributes x1a, x3a and x4a), RP design does not perform significantly better that ME design. However, results from model (5) indicate that RD design is significantly better than ME design in WTP estimates of attributes x1a and x3a.

Conclusion

The increasing application of choice experiments in consumer valuation reflects the comparative advantage of this technique in valuation of multiple product attributes and the tradeoff between different attributes can be easily estimated in a single experiment rather than multiple surveys or experiments required by other valuation techniques. On the other hand, the popularity of choice experiments also demands more research on this area to study the impacts of design parameters on consumer valuation estimates. Although used by many researchers, the effort to address the second issues is not adequate. Carlsson and Martinsson (2003), Lusk and Norwood (2005), and Ferrini and Scarpa (2007) are among the few to study the impacts of design strategies on consumer valuation estimate using Monte Carlo simulation. However, in those studies, the impacts of choice experiment design are evaluated with fixed attribute information loads. With the current field studies finding that attribute information loads affect consumer WTP estimates, it is worthwhile to investigate the performance of different design strategies under different information loads, as well as to identify the sources that affect consumer valuation in field study—are the impacts of information load due to the change in consumer cognitive ability to make choice decision, the substitute and complement effects between product attributes or the pure statistical property of choice experiments? In this paper we extend previous researches by investigating the change in consumer WTP estimates under different attribute information loads, design strategies, sample sizes and utility function types to address those questions.

Results in this article delivered similar information with previous studies that larger sample size can significantly improve the WTP estimates. If all other design parameters such as information loads and sample size are controlled, RD design is the overall best design; however, this design is not significantly better than RP design. If other design parameters are also considered, the information loads significantly affect the performance of design strategies — when the number of attributes is 5, RD design is the best with both continuous and discrete cases; when the number of attributes is 3, RP and MD are the best with both continuous and discrete utility functions, respectively. The fact that the WTP estimates have a quadratic relationship with the number of attributes in choice experiments is unexpected. This result calls for further research to investigate the impacts of attribute information on consumer valuation—how much of the changes in consumer valuation found by the previous researches such as Hensher (2006) and Gao and Schroder (2009) result from the changes in consumer cognitive ability to processing information, or due to the statistical property of the choice experiments with the change of the number of attributes in a design? The quadratic relationships between the performance and the number of attributes in choice experiments also indicates that keeping the information loads as few as possible is not the best design strategies to improve consumer valuation estimates. However, too many attributes will deteriorate the performance of all design strategies in valuation estimates. In our case, the optimal attribute number in the choice experiment was 3 in the discrete case. However, this conclusion depends on the particular attributes studied.

Choosing right design strategies is only a small step for a successful choice experiment. Many factors are correlated so that tradeoffs must be made between statistical performance, feasibility to administer the choice experiments, the potential problem of omitted variables in a consumer utility function, and the budget constraints of the researchers. For example, the number of attributes in a choice experiment determines the minimum choice set in ME and MD designs. ME design may be an overall better design than MD design, but the choice sets in ME design is always larger than that in MD design. The increased number of choice sets in an ME design may increase the cognitive burden on respondents, which may result in less accurate estimation. Restricting information loads in a choice experiment may reduce cognitive burden on respondents and improve the statistical property of choice experiments. However, if the omitted information includes important variables determining consumer preference in real-world purchases, less attribute information may result in biased estimation of consumer preference. RD design overall perform the best. However, it is not financially efficient to ask one respondent to only answer one choice experiment question. All those factors should be carefully evaluated before conducting a choice experiment. Further research may add more design parameters such as the number of choices in choice sets in choice experiments to study the performance of various design strategies. Furthermore, worthwhile is expressing the deviation of WTP estimates from the true WTP as a function of major factors that affect consumer valuation subject to research budget, and seek the optimal strategies regarding experiment design, attribute information loads, sample size, and other design parameters.

[1] Because each attribute has three levels, two dummy variables are created for each attribute.

References:

Adamowicz, W., R. Boxall, M. Williams, and J. Louviere. 1998. “Stated Preference Approaches for Measuring Passive Use values: Choice Experiments and Contingent Valuation.” *American Journal of Agricultural Economics* 80:64-75.

Alfnes, F. 2004. “Stated Preferences for Imported and Hormone-Treated Beef: Application of a Mixed Logit Model.” *European Review of Agricultural Economics* 31:19-37.

Boyle, K.J., T.P. Holmes, M.F. Teisl, and B. Roe. 2001. “A Comparison of Conjoint Analysis Response Formats.” American *Journal of Agricultural Economics* 83:441-54.

Carlsson, F., and P. Martinsson. 2003. “Design Techniques for Stated Preference Methods in Health Economics.” *Health Economics* 12:281-94.

Chrzan,K., and B. Orme. “An Overview and Comparison of Design Strategies for Choice-Based Conjoint Analysis.” Sawtooth Software Research Paper Series, Sawtooth Software, Sequim, WA, 2000.

Colombo, S., J. Calatrava-Requena, and N. Hanley. 2007. “Testing Choice Experiment for Benefit Transfer with Preference Heterogeneity.” *American Journal of Agricultural Economics* 2007: 135-51.

Darby, K, M. T. Batte, S. Ernst and B. Roe. 2008. “Decomposing Local: A Conjoint Analysis of Locally Produced Foods.” *American Journal of Agricultural Economics* 90: 476-86.

Ferrini, S. and R. Scarpa. 2007. “Designs with a priori Information for Nonmarket Valuation with Choice Experiments: A Monte Carlo Study.” *Journal of Environmental Economics and Management* 53:342-63.

Gao, Z. and T.C. Schroder. 2009. “Effects of Label Information on Consumer Willingness-To-Pay for Food Attributes.” *American Journal of Agricultural Economics forthcoming*.

Hausmand, J.A., ed. Contingent Valuation: A Critical Assessment. Amsterdam: North Holland, 1993.

Hensher, D.A. 2006. "Revealing Difference in Willingness to Pay due to the Dimensionality of Stated Choice Designs: An Initial Assessment." *Environmental and Resource Economics* 34:7–44.

Islam, T., J.J. Louviere, and P.F. Burke. 2007. "Modeling the Effects of Including/Excluding Attributes in Choice Experiments on Systematic and Random Components." *International Journal of Research in Marketing* 24:289–300.

Lusk, J.L., and F.B. Norwood. 2005. "Effect of Experimental Design on Choice-Based Conjoint Valuation Estimates." American Journal of Agricultural Economics 87:771-85.

McFadden D. 1974. +. In *Frontiers in Econometrics,* Zarembka P (ed.) Academic Press: New York, 105-142.

Tonsor, G.T., T.C. Schroeder, J.A. Fox and A. Biere. 2005. "European Preferences for Beef Steak Attributes." *Journal of Agricultural and Resource Economics* 30:367-380.

Table 1. Description of Simulation Scenarios

Experiment Design	Description	# of Attribute (N)	# of Choice Sets in Experiment	Utility Function	Sample Size
RD	Random drew choice set from N^3xN^3 full factorial design	2	N/A	Continuous /Discrete	243/486/972
		3	N/A		250/500/1000
		4	N/A		250/500/1000
		5	N/A		250/500/1000
ME	Main effects only design drawn from N^3xN^3 full factorial design	2	9	Continuous /Discrete	252/504/1008
		3	27		270/513/1026
		4	27		270/513/1026
		5	27		270/513/1026
MD	Main effects minimum design with maximized D-efficiency	2	9	Continuous /Discrete	252/504/1008
		3	13		260/507/1001
		4	17		255/510/1003
		5	21		252/504/1008
RP	Random pair alternatives generated from full factorial design	2	9	Continuous /Discrete	243/495/999
		3	27		243/486/999
		4	81		243/486/972
		5	243		243/486/972

Table 2 WTP Estimates, MSE, RSE and Best Design under Different Simulation Scenarios (Continuous Utility Function)

Attribute	# of Attribute	RD			ME			MD			RP			Best Design		
		s=S	s=M	s=L	s=S	s=M	s=L	s=S	s=M	s=L	s=S	s=M	s=L	s=S	s=M	s=L
x1	2	2.17[a]	2.05	2.03	2.03	2.03	2.03	2.15	2.08	2.01	2.29	2.18	2.08			
		(0.53)[b]	(0.17)	(0.87)	(0.17)	(0.10)	(0.50)	(0.47)	(0.21)	(0.08)	(1.23)	(0.43)	(0.15)	ME[d]	ME	MD
		{0.25}[c]	{0.16}	{0.12}	{0.17}	{0.13}	{0.09}	{0.26}	{0.18}	{0.12}	{0.37}	{0.25}	{0.15}	ME[e]	ME	ME
	3	2.25	2.09	2.04	2.35	2.21	2.11	2.20	2.16	2.11	2.20	2.08	2.02			
		(0.86)	(0.28)	(0.12)	(1.78)	(0.71)	(0.25)	(0.90)	(0.47)	(0.25)	(0.80)	(0.27)	(0.11)	RP	RP	RP
		{0.33}	{0.20}	{0.14}	{0.43}	{0.28}	{0.19}	{0.35}	{0.26}	{0.19}	{0.32}	{0.20}	{0.13}	RP	RP	RP
	4	2.13	2.14	2.03	3.38	2.38	2.13	1.50	1.80	2.22	2.22	2.18	2.07			
		(0.74)	(0.46)	(0.17)	(7.67)	(1.30)	(0.17)	(5.14)	(2.72)	(1.29)	(0.89)	(0.53)	(0.20)	RD	RD	RD
		{0.33}	{0.26}	{0.16}	{0.85}	{0.34}	{0.16}	{0.80}	{0.55}	{0.38}	{0.35}	{0.27}	{0.18}	RD	RD	RD
	5	2.41	2.20	2.10	2.33	2.22	2.12	1.20	2.16	2.24	2.45	2.21	2.09			
		(2.37)	(0.55)	(0.26)	(2.12)	(0.67)	(0.26)	(6.65)	(4.21)	(0.95)	(2.34)	(0.68)	(0.25)	ME	RD	RP
		{0.47}	{0.28}	{0.19}	{0.48}	{0.30}	{0.20}	{0.83}	{0.61}	{0.35}	{0.50}	{0.30}	{0.19}	RD	RD	RP
x2	3	3.40	3.14	3.07	3.52	3.30	3.16	3.33	3.26	3.19	3.29	3.12	3.03			
		(2.00)	(0.62)	(0.25)	(3.62)	(1.40)	(0.48)	(2.54)	(1.35)	(0.72)	(1.61)	(0.52)	(0.22)	RP	RP	RP
		{0.33}	{0.21}	{0.14}	{0.41}	{0.27}	{0.18}	{0.40}	{0.29}	{0.22}	{0.29}	{0.19}	{0.13}	RP	RP	RP
	4	3.20	3.21	3.05	4.49	3.46	3.17	1.88	2.62	3.41	3.36	3.27	3.12			
		(1.60)	(1.00)	(0.36)	(9.03)	(1.80)	(0.32)	(12.89)	(8.36)	(4.41)	(2.05)	(1.08)	(0.41)	RD	RD	ME
		{0.32}	{0.24}	{0.16}	{0.65}	{0.29}	{0.14}	{0.91}	{0.67}	{0.47}	{0.34}	{0.25}	{0.17}	RD	RD	ME
	5	3.57	3.28	3.14	3.57	3.36	3.19	1.60	3.25	3.40	3.63	3.28	3.14			
		(4.45)	(1.09)	(0.50)	(5.61)	(1.65)	(0.63)	(19.30)	(12.36)	(2.69)	(4.87)	(1.48)	(0.57)	RD	RD	RD
		{0.43}	{0.26}	{0.18}	{0.51}	{0.31}	{0.21}	{0.94}	{0.68}	{0.39}	{0.49}	{0.30}	{0.19}	RD	RD	RD
x3	4	4.27	4.27	4.05	6.75	4.76	4.25	2.56	3.46	4.57	4.50	4.38	4.16			
		(2.95)	(1.79)	(0.68)	(30.76)	(5.16)	(0.65)	(28.56)	(15.90)	(8.28)	(3.86)	(2.00)	(0.75)	RD	RD	ME
		{0.33}	{0.25}	{0.17}	{0.85}	{0.33}	{0.15}	{0.97}	{0.68}	{0.48}	{0.35}	{0.26}	{0.17}	RD	RD	ME
	5	4.81	4.39	4.20	4.77	4.49	4.26	2.61	4.28	4.44	4.89	4.39	4.19			
		(8.56)	(2.02)	(0.92)	(9.74)	(2.90)	(1.08)	(19.95)	(13.40)	(3.18)	(9.33)	(2.61)	(1.02)	RD	RD	RD
		{0.45}	{0.27}	{0.18}	{0.51}	{0.31}	{0.20}	{0.74}	{0.55}	{0.32}	{0.50}	{0.30}	{0.20}	RD	RD	RD
x4	5	0.11	0.12	0.11	0.09	0.10	0.11	0.17	0.07	0.08	0.11	0.10	0.10			
		(0.05)	(0.02)	(0.01)	(0.07)	(0.02)	(0.01)	(0.13)	(0.06)	(0.01)	(0.06)	(0.02)	(0.01)	RD	RP	RD
		{1.78}	{1.18}	{0.85}	{2.01}	{1.19}	{0.83}	{2.36}	{1.64}	{0.88}	{1.91}	{1.13}	{0.81}	RD	RP	RP

Notes: We removed 5% of the lowest and highest extreme value in WTP estimates, because some of those values are not rescannable.
[a] Mean WTP from 450 simulated WTP
[b] Mean Squared Errors
[c] Average of the Absolute Relative Error
[d] Best Design based on Minimum MSE
[e] Best Design based on Minimum RSE

Table 3 WTP Estimates, MSE, RSE and Best Design under Different Simulation Scenarios (Discrete Utility Function)

Attribute	# of Attribute	RD			ME			MD			RP			Best Design		
		s=S	s=M	s=L	s=S	s=M	s=L	s=S	s=M	s=L	s=S	s=M	s=L	s=S	s=M	s=L
x1a	2	2.12	2.03	2.00	2.04	2.02	2.03	2.08	2.03	2.00	2.10	2.04	2.00			
		(0.37)	(0.13)	(0.06)	(0.29)	(0.13)	(0.07)	(0.25)	(0.11)	(0.05)	(0.45)	(0.19)	(0.07)	MD	MD	MD
		{0.24}	{0.15}	{0.10}	{0.21}	{0.14}	{0.11}	{0.20}	{0.13}	{0.09}	{0.26}	{0.18}	{0.11}	MD	MD	MD
	3	2.10	2.08	2.04	2.19	2.08	2.04	2.11	2.05	2.03	2.13	2.07	2.03			
		(0.44)	(0.21)	(0.10)	(0.66)	(0.25)	(0.11)	(0.40)	(0.18)	(0.09)	(0.41)	(0.20)	(0.08)	MD	MD	RP
		{0.26}	{0.18}	{0.13}	{0.30}	{0.19}	{0.14}	{0.24}	{0.17}	{0.12}	{0.25}	{0.18}	{0.11}	MD	MD	RP
	4	2.16	2.08	2.02	2.11	2.07	2.06	2.34	2.24	2.15	2.20	2.09	2.06			
		(0.61)	(0.24)	(0.09)	(0.51)	(0.23)	(0.12)	(2.44)	(0.88)	(0.37)	(0.76)	(0.28)	(0.14)	ME	ME	RD
		{0.30}	{0.20}	{0.13}	{0.28}	{0.19}	{0.14}	{0.53}	{0.35}	{0.23}	{0.32}	{0.20}	{0.15}	ME	ME	RD
	5	2.27	2.11	2.06	2.32	2.23	2.10	1.28	2.00	2.21	2.27	2.18	2.10			
		(1.03)	(0.28)	(0.13)	(1.85)	(0.77)	(0.26)	(4.90)	(1.73)	(0.79)	(1.30)	(0.54)	(0.24)	RD	RD	RD
		{0.37}	{0.20}	{0.14}	{0.48}	{0.31}	{0.20}	{0.80}	{0.47}	{0.32}	{0.40}	{0.27}	{0.19}	RD	RD	RD
x1b	2	1.06	1.01	0.99	1.02	1.01	1.02	1.05	1.02	1.01	1.05	1.02	1.00			
		(0.14)	(0.05)	(0.02)	(0.09)	(0.04)	(0.02)	(0.09)	(0.04)	(0.02)	(0.09)	(0.04)	(0.02)	ME	RP	RP
		{0.29}	{0.18}	{0.13}	{0.24}	{0.17}	{0.12}	{0.24}	{0.17}	{0.12}	{0.23}	{0.16}	{0.11}	RP	RP	RP
	3	1.05	1.05	1.02	1.08	1.04	1.03	1.07	1.04	1.02	1.06	1.03	1.01			
		(0.17)	(0.08)	(0.04)	(0.18)	(0.07)	(0.03)	(0.14)	(0.06)	(0.03)	(0.22)	(0.10)	(0.04)	MD	MD	MD
		{0.33}	{0.24}	{0.16}	{0.32}	{0.21}	{0.15}	{0.30}	{0.19}	{0.14}	{0.37}	{0.25}	{0.15}	MD	MD	MD
	4	1.06	1.03	1.01	1.05	1.03	1.02	1.16	1.10	1.06	1.12	1.05	1.04			
		(0.18)	(0.08)	(0.04)	(0.23)	(0.11)	(0.06)	(0.58)	(0.19)	(0.08)	(0.32)	(0.12)	(0.05)	RD	RD	RD
		{0.35}	{0.23}	{0.16}	{0.38}	{0.27}	{0.20}	{0.51}	{0.33}	{0.23}	{0.42}	{0.27}	{0.19}	RD	RD	RD
	5	1.16	1.06	1.03	1.18	1.10	1.06	0.89	0.99	1.04	1.11	1.10	1.06			
		(0.43)	(0.11)	(0.06)	(0.60)	(0.22)	(0.10)	(1.79)	(0.56)	(0.27)	(0.45)	(0.20)	(0.10)	RD	RD	RD
		{0.48}	{0.27}	{0.19}	{0.56}	{0.36}	{0.24}	{0.94}	{0.59}	{0.43}	{0.49}	{0.34}	{0.24}	RP	RD	RD
x2a	3	3.15	3.12	3.06	3.30	3.15	3.06	3.17	3.08	3.04	3.25	3.13	3.03			
		(0.79)	(0.39)	(0.19)	(1.49)	(0.57)	(0.27)	(0.88)	(0.34)	(0.17)	(0.97)	(0.49)	(0.18)	RD	MD	MD
		{0.23}	{0.17}	{0.12}	{0.30}	{0.20}	{0.14}	{0.23}	{0.16}	{0.11}	{0.26}	{0.19}	{0.11}	RD	MD	MD
	4	3.25	3.14	3.04	3.12	3.08	3.09	3.41	3.28	3.17	3.30	3.13	3.09			
		(1.28)	(0.52)	(0.22)	(1.10)	(0.49)	(0.28)	(2.99)	(1.10)	(0.45)	(1.54)	(0.54)	(0.26)	ME	ME	RD
		{0.28}	{0.19}	{0.13}	{0.27}	{0.19}	{0.14}	{0.39}	{0.25}	{0.17}	{0.31}	{0.19}	{0.13}	ME	ME	RD
	5	3.43	3.19	3.10	3.40	3.27	3.13	1.90	2.89	3.21	3.35	3.25	3.14			

		(2.70)	(0.79)	(0.35)	(3.48)	(1.27)	(0.49)	(11.12)	(4.02)	(1.73)	(2.08)	(0.97)	(0.12)	RP	RD	RP
		{0.39}	{0.22}	{0.16}	{0.42}	{0.27}	{0.18}	{0.83}	{0.49}	{0.34}	{0.34}	{0.24}	{0.17}	RP	RD	RD
	3	2.11	2.06	2.04	2.19	2.09	2.05	2.12	2.05	2.02	2.13	2.07	2.02			
		(0.38)	(0.17)	(0.08)	(0.64)	(0.26)	(0.12)	(0.37)	(0.13)	(0.07)	(0.37)	(0.19)	(0.08)	MD	MD	MD
		{0.24}	{0.17}	{0.12}	{0.30}	{0.20}	{0.14}	{0.22}	{0.14}	{0.10}	{0.24}	{0.17}	{0.11}	MD	MD	MD
x2b	4	2.16	2.09	2.03	2.13	2.08	2.07	2.30	2.21	2.12	2.17	2.07	2.05			
		(0.57)	(0.23)	(0.10)	(0.49)	(0.22)	(0.11)	(1.89)	(0.65)	(0.27)	(0.57)	(0.21)	(0.11)	ME	RP	RD
		{0.28}	{0.19}	{0.13}	{0.28}	{0.19}	{0.13}	{0.46}	{0.29}	{0.19}	{0.29}	{0.18}	{0.13}	ME	RP	RD
	5	2.28	2.13	2.07	2.25	2.20	2.09	1.01	2.07	2.24	2.24	2.17	2.09			
		(1.35)	(0.40)	(0.17)	(1.59)	(0.66)	(0.20)	(7.55)	(2.55)	(1.07)	(1.02)	(0.46)	(0.20)	RP	RD	RD
		{0.41}	{0.24}	{0.16}	{0.44}	{0.28}	{0.13}	{0.94}	{0.56}	{0.38}	{0.36}	{0.25}	{0.18}	RP	RD	ME
x3a	4	4.31	4.18	4.04	4.16	4.08	4.09	4.51	4.34	4.21	4.44	4.19	4.13			
		(2.06)	(0.89)	(0.36)	(1.62)	(0.67)	(0.39)	(4.71)	(1.72)	(0.69)	(3.38)	(1.12)	(0.51)	ME	ME	RD
		{0.26}	{0.18}	{0.13}	{0.24}	{0.17}	{0.12}	{0.37}	{0.24}	{0.16}	{0.33}	{0.20}	{0.14}	ME	ME	ME
	5	4.56	4.26	4.15	4.43	4.36	4.15	2.70	4.16	4.44	4.50	4.35	4.20			
		(4.41)	(1.26)	(0.59)	(6.54)	(2.70)	(0.86)	(27.12)	(8.78)	(3.02)	(4.14)	(1.76)	(0.79)	RP	RD	RD
		{0.37}	{0.22}	{0.15}	{0.44}	{0.29}	{0.18}	{0.87}	{0.48}	{0.31}	{0.35}	{0.24}	{0.17}	RP	RD	RD
x3b	4	3.22	3.13	3.03	3.08	3.04	3.07	3.39	3.29	3.17	3.35	3.14	3.10			
		(1.15)	(0.52)	(0.22)	(0.97)	(0.44)	(0.25)	(3.16)	(1.25)	(0.48)	(1.96)	(0.65)	(0.29)	ME	ME	RD
		{0.27}	{0.19}	{0.13}	{0.25}	{0.18}	{0.13}	{0.40}	{0.27}	{0.17}	{0.34}	{0.21}	{0.14}	ME	ME	RD
	5	3.44	3.20	3.09	3.35	3.29	3.13	1.98	3.01	3.29	3.36	3.25	3.14			
		(2.56)	(0.73)	(0.34)	(3.34)	(1.41)	(0.45)	(14.18)	(4.76)	(1.84)	(2.18)	(0.89)	(0.42)	RP	RD	RD
		{0.38}	{0.22}	{0.15}	{0.43}	{0.28}	{0.18}	{0.88}	{0.50}	{0.34}	{0.34}	{0.23}	{0.16}	RP	RD	RD
x4a	5	5.74	5.34	5.18	5.72	5.51	5.20	3.38	5.05	5.43	5.66	5.47	5.23			
		(7.06)	(2.06)	(0.92)	(10.44)	(4.06)	(1.19)	(27.89)	(9.57)	(3.71)	(6.90)	(3.12)	(1.28)	RP	RD	RD
		{0.37}	{0.22}	{0.15}	{0.43}	{0.28}	{0.17}	{0.73}	{0.41}	{0.28}	{0.36}	{0.26}	{0.17}	RP	RD	RD
x4b	5	4.59	4.28	4.15	4.64	4.40	4.17	2.82	4.00	4.33	4.52	4.38	4.18			
		(4.35)	(1.32)	(0.58)	(8.11)	(2.77)	(0.94)	(15.50)	(5.61)	(2.25)	(4.53)	(1.94)	(0.83)	RD	RD	RD
		{0.38}	{0.22}	{0.15}	{0.47}	{0.30}	{0.19}	{0.70}	{0.40}	{0.27}	{0.36}	{0.25}	{0.17}	RD	RD	RD

Notes: We removed 5% of the lowest and highest extreme value in WTP estimates, because some of those values are not rescannable.

[a] Mean WTP from 450 simulated WTP

[b] Mean Squared Errors

[c] Average of the Absolute Relative Error

[d] Best Design based on Minimum MSE

[e] Best Design based on Minimum RSE

Table 4 Models of WTP Estimates, SE, RE on Design Strategy, Samples Size and Number of Attributes (Continuous Utility Function)

	WTP for				SE for				RE for			
	X1	X2	X3	X4	X1	X2	X3	X4	X1	X2	X3	X4
Design Strategy												
RD	0.15	0.35	0.68	0.00	-1.40	-5.86	-12.06	-0.04	-0.17	-0.30	-0.35	-0.36
	(0.00)[a]	(0.00)	(0.00)	(0.70)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ME	0.29	0.59	1.23	-0.01	-0.68	-4.45	-6.50	-0.03	-0.10	-0.22	-0.23	-0.28
	(0.00)	(0.00)	(0.00)	(0.33)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
RP	0.19	0.37	0.77	0.00	-1.29	-5.76	-11.62	-0.04	-0.14	-0.29	-0.33	-0.35
	(0.00)	(0.00)	(0.00)	(0.61)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Sample Size												
Middle	-0.07	-0.02	-0.09	-0.03	-1.31	-3.07	-8.49	-0.05	-0.16	-0.17	-0.22	-0.73
	(0.00)	(0.50)	(0.14)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Large	-0.11	-0.06	-0.13	-0.02	-1.88	-4.83	-12.14	-0.07	-0.26	-0.29	-0.35	-1.17
	(0.00)	(0.06)	(0.03)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
# of Attributes	0.19	-0.27	-0.02	N/A[d]	1.02	7.02	-2.22	N/A	0.18	0.50	-0.04	N/A
	(0.00)	(0.25)	(0.66)	N/A	(0.00)	(0.00)	(0.00)	N/A	(0.00)	(0.00)	(0.00)	N/A
# of Attributes Squared	-0.02	0.03	N/A[c]	N/A	-0.07	-0.67	N/A	N/A	-0.02	-0.05	N/A	N/A
	(0.00)	(0.28)	N/A	N/A	(0.04)	(0.00)	(0.00)	N/A	(0.00)	(0.00)	(0.00)	N/A
Constant	1.71	3.47	3.83	0.12	0.33	-7.10	31.74	0.10	0.33	-0.41	0.98	2.26
	(0.00)	(0.00)	(0.00)	(0.00)	(0.39)	(0.03)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Statistical Test												
RD=ME	(0.00)	(0.00)	(0.00)	(0.17)	(0.00)	(0.00)	(0.00)	(0.17)	(0.00)	(0.00)	(0.00)	(0.16)
RD=RP	(0.08)	(0.61)	(0.24)	(0.37)	(0.23)	(0.71)	(0.52)	(0.60)	(0.00)	(0.37)	(0.14)	(0.86)
ME=RP	(0.00)	(0.00)	(0.00)	(0.64)	(0.00)	(0.00)	(0.00)	(0.40)	(0.00)	(0.00)	(0.00)	(0.22)
Middle=Large	(0.01)	(0.23)	(0.52)	(0.75)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R-Squared	0.013	0.015	0.027	0.004	0.060	0.068	0.075	0.093	0.122	0.144	0.134	0.128
# Observations	21600[b]	16200	10800	5400	21600	16200	10800	5400	21600	16200	10800	5400

[a] Numbers in parentheses are *p* values.
[b] Number of observations are derived from 450 simulations from each simulation scenario.
[c] On attribute X3, the # of attributes squared variable is nearly perfectly collinear with the number of attributes variable.
[d] On attribute X4, there is no variation in the # of attribute, only choice experiments with five attributes include X4.

Table 5 Models of WTP Estimates, SE, RE on Design Strategy, Samples Size and Number of Attributes (Discrete Utility Function)

	WTP for				SE for				RE for			
	X1a	X2a	X3a	X4a	X1a	X2a	X3a	X4a	X1a	X2a	X3a	X4a
Design Strategy												
RD	0.05	0.15	-0.15	0.80	-0.71	-1.73	-6.08	-10.38	-0.10	-0.12	-0.19	-0.22
	(0.00)[a]	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ME	0.06	0.16	-0.13	0.85	-0.58	-1.48	-5.54	-8.50	-0.08	-0.09	-0.16	-0.18
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
RP	0.06	0.17	-0.16	0.83	-0.63	-1.71	-5.72	-9.96	-0.09	-0.11	-0.16	-0.21
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Sample Size												
Middle	-0.02	-0.03	-0.13	0.22	-0.65	-1.58	-4.39	-8.37	-0.12	-0.12	-0.15	-0.18
	(0.04)	(0.19)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Large	-0.06	-0.07	-0.21	0.14	-0.87	-2.12	-5.85	-11.30	-0.19	-0.20	-0.23	-0.28
	(0.00)	(0.00)	(0.00)	(0.10)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
# of Attributes	0.15	0.46	-0.26	N/A[d]	-0.58	-3.93	3.66	N/A	-0.05	-0.26	0.12	N/A
	(0.00)	(0.00)	(0.00)	N/A	(0.00)	(0.00)	(0.00)	N/A	(0.00)	(0.00)	(0.00)	N/A
# of Attributes Squared	-0.02	-0.06	N/A[c]	N/A	0.13	0.61	N/A	N/A	0.02	0.04	N/A	N/A
	(0.00)	(0.00)	N/A	N/A	(0.00)	(0.00)	N/A	N/A	(0.00)	(0.00)	N/A	N/A
Statistical Test												
RD=ME	(0.20)	(0.58)	(0.45)	(0.55)	(0.00)	(0.02)	(0.19)	(0.03)	(0.00)	(0.00)	(0.01)	(0.00)
RD=RP	(0.26)	(0.43)	(0.28)	(0.70)	(0.06)	(0.81)	(0.38)	(0.63)	(0.00)	(0.34)	(0.02)	(0.31)
ME=RP	(0.87)	(0.81)	(0.07)	(0.84)	(0.24)	(0.03)	(0.65)	(0.09)	(0.17)	(0.00)	(0.80)	(0.02)
Middle=Large	(0.01)	(0.04)	(0.14)	(0.33)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Constant	1.79	2.19	0.78	4.50	1.82	9.32	-5.37	20.28	0.37	0.77	-0.03	0.63
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.31)	(0.00)
R-Squared	0.004	0.005	0.003	0.004	0.071	0.083	0.067	0.075	0.163	0.144	0.142	0.137
# Observations	21600[b]	16200	10800	5400	21600	16200	10800	5400	21600	16200	10800	5400

Notes: The models for attributes Xib (i=1,2,3,4) are not presented in the table because the results are similar with those from models for attributes Xia.

[a] Numbers in parentheses are *p* values.

[b] Number of observations are derived from 450 simulations from each simulation scenario.

[c] On attribute X3a, the # of attributes squared variable is nearly perfectly collinear with the number of attributes variable.

[d] On attribute X4a, there is no variation in the # of attribute, only choice experiments with five attributes include X4a.